

基于改进 YOLOv8 的脑肿瘤检测模型

张傲, 刘微, 刘阳, 杨思瑶, 管勇, 李波, 刘芳菲

(沈阳理工大学 信息科学与工程学院, 沈阳 110159)

摘要: 为应对磁共振成像(MRI)中形态复杂和边界不规则的脑肿瘤检测, 提出一种改进的 YOLOv8 模型 TumorNet-YOLO。该模型通过三项创新模块提升检测性能: 自适应感受野卷积模块增强多尺度肿瘤特征的提取能力, 降低漏检率; 分割融合卷积模块通过多尺度特征融合, 增强浅层和深层特征的协同作用; 可变形融合模块优化对不规则肿瘤区域的检测, 提升模型在复杂 MRI 背景下的鲁棒性。实验结果表明, TumorNet-YOLO 在脑肿瘤检测数据集 Br35H 中表现优异, 平均精度均值 $mAP@0.5$ 为 96.6%, $mAP@0.5:0.95$ 为 73.8%。此外, 模型计算量 (GFLOPs) 为 8.6, 显著优于现有方法。为了验证模型的泛化能力, 在 BCCD 和 BTOD 数据集上进行了对比实验, 结果显示 TumorNet-YOLO 在 $mAP@0.5$ 和 $mAP@0.5:0.95$ 等多个指标上超越了 YOLOv8n, 表明 TumorNet-YOLO 可为脑肿瘤检测和医学图像分析提供更为有效的解决方案。

关键词: 脑肿瘤检测; 医学影像分析; YOLOv8; 自适应感受野

中图分类号: TP391.41 文献标志码: A DOI:10.3969/j.issn.1003-1251.2025.05.002

Brain Tumor Detection Model Based on Improved YOLOv8

ZHANG Ao, LIU Wei, LIU Yang, YANG Siyao, GUAN Yong, LI Bo, LIU Fangfei

(Shenyang Ligong University, Shenyang 110159, China)

Abstract: In order to address the challenges of detecting brain tumors with complex shapes and irregular boundaries in Magnetic Resonance Imaging (MRI), an improved YOLOv8 model, TumorNet-YOLO, is proposed. This model enhances detection performance through three innovative modules: the Adaptive Receptive Field Convolution Module improves the ability to extract multi-scale tumor features and reduce false negatives; the Segmentation Fusion Convolution Module strengthens the synergy between shallow and deep features through multi-scale feature fusion; the Deformable Fusion Module optimizes the detection of irregular tumor regions and improves the model's robustness in complex MRI backgrounds. Experimental results show that TumorNet-YOLO performs excellently on the brain tumor detection dataset Br35H, with a mean Average Precision $mAP@0.5$ of 96.6% and $mAP@0.5:0.95$ of 73.8%. Furthermore, the model's computational cost is 8.6 GFLOPs, which significantly outperforms existing methods. To validate the model's generalization ability, comparative experiments were conducted on the BCCD and BTOD datasets. The results demonstrate that TumorNet-YOLO outperforms YOLOv8n in multiple metrics, including $mAP@0.5$ and $mAP@0.5:0.95$, indicating that TumorNet-YOLO provides an effective solution for brain tumor detection and medical image analysis.

收稿日期: 2024-12-27

基金项目: 辽宁省教育厅高等学校基本科研项目(JYTMS20230189); 沈阳理工大学引进高层次人才科研支持计划项目(1010147001131)

作者简介: 张傲(2000—), 男, 硕士研究生; 刘微(1978—), 通信作者, 女, 教授。

Key words: brain tumor detection; medical image analysis; YOLOv8; adaptive receptive field

医学影像分析是脑肿瘤诊断的关键环节。磁共振成像 (MRI) 作为一种重要的软组织检测工具,能够在无创条件下准确展示脑肿瘤的位置、大小和形态。然而,传统的人工分析方法效率低且容易受到主观因素的影响,无法满足临床对准确性和效率的要求。近年来,基于卷积神经网络 (CNN) 的自动化检测方法取得了显著进展,并已广泛应用于医学图像分析。YOLO 系列模型因其高效且精确的目标检测能力,逐渐成为研究人员解决脑肿瘤检测问题的重要工具。然而,标准 YOLO 模型在处理背景复杂以及病变区域不规则的医学图像时存在一定的局限性。

随着机器学习和深度学习技术的发展,研究者们提出了多种方法以提高脑肿瘤检测的准确性和效率。Ranjbarzadeh 等^[1]利用深度卷积神经网络 (DCNN) 对 MRI 图像中的脑肿瘤进行检测和分类,证明了 DCNN 在处理复杂病变区域时的有效性。Chattopadhyay 等^[2]通过优化 CNN 模型的特征提取和数据增强策略,显著提高了检测精度。Tahir 等^[3]提出的多模态融合深度学习模型结合了多视图 CNN 架构,能够更好地捕捉肿瘤的空间特征。与此同时,强化学习与 CNN 的结合也展现了显著的潜力,能够快速高效地进行肿瘤的自动识别和定位。

YOLO 系列模型及其改进版本在脑肿瘤检测中表现出了卓越的性能。YOLOv5 在脑癌分割任务中具备的准确度和敏感性,证明了其在医学图像分析中的能力^[4]。RCS-YOLO^[5]专为脑肿瘤检测设计,不仅提高了检测速度,还增强了精度,且在脑肿瘤检测数据集 Br35H 上的表现超过了 YOLOv6、YOLOv7 和 YOLOv8,具有更高的速度和准确性^[6],即便 YOLOv3 在脑癌检测和定位中的应用也在精度和召回率方面表现出色,证明了 YOLO 系列在脑肿瘤检测中的可靠性^[7]。与 Faster R-CNN、SSD 等其他深度学习模型相比,YOLO 系列在处理复杂病变区域时具备更快的检测速度,同时保持较高的精度^[8]。尽管现有的 YOLO 系列模型在检测精度和速度上取得了显著进展,仍面临一些挑战。首先,YOLO 模型在处理 MRI 图像中的复杂背景时,存在对背景和肿瘤之间界限不明确的情况,难以有效区分肿瘤与周围组织,导致漏检或误检的发生。其次,YOLO 系列模型依赖矩形框进行目标定位,而脑肿瘤通常呈

现不规则形态,会导致传统的矩形框无法精确地包围肿瘤区域,从而限制了模型对形态复杂肿瘤的检测。

为了克服上述问题,本研究提出了一种基于 YOLOv8 的改进模型 TumorNet-YOLO。该模型引入了自适应感受野卷积 (ARFConv) 模块,通过动态调整感受野的大小,以适应不同尺度的病变特征;设计了分割融合卷积模块 (SFCM),优化多尺度特征融合以降低计算复杂度;采用了可变形融合模块 (DFB),增强对复杂形态病变区域的检测能力。

1 TumorNet-YOLO 模型

TumorNet-YOLO 是专为脑肿瘤 MRI 图像检测设计的改进模型,通过优化特征提取和融合能力,有效应对不规则脑肿瘤区域的检测挑战。基于 YOLOv8 的架构,TumorNet-YOLO 整体架构如图 1 所示,图中 CBS 表示由卷积 (Conv2d)、批归一化 (BatchNorm2d) 和激活函数 (SiLU) 组成的组合卷积层。

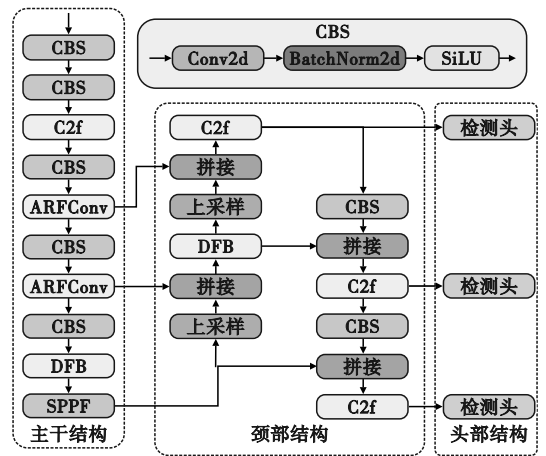


图 1 TumorNet-YOLO 整体架构

Fig. 1 Overall architecture of TumorNet-YOLO

1.1 自适应感受野卷积模块

在计算机视觉中,传统卷积神经网络在脑肿瘤检测中面临两大问题:固定感受野难以捕捉多尺度特征,且卷积核的参数共享机制不足以处理肿瘤的复杂形态和局部差异。MRI 图像中的肿瘤通常具有多尺度和不规则形态,且背景复杂,增加了检测的难度。为此,本文提出 ARFConv 模块,通过动态调整感受野的大小和形状,使模型能够

自适应不同尺度的肿瘤特征,并结合通道注意力机制,提升特征选择的准确性和灵活性,其结构如图2所示。ARFConv能够有效抑制背景噪声,并精准捕捉肿瘤区域的关键特征。

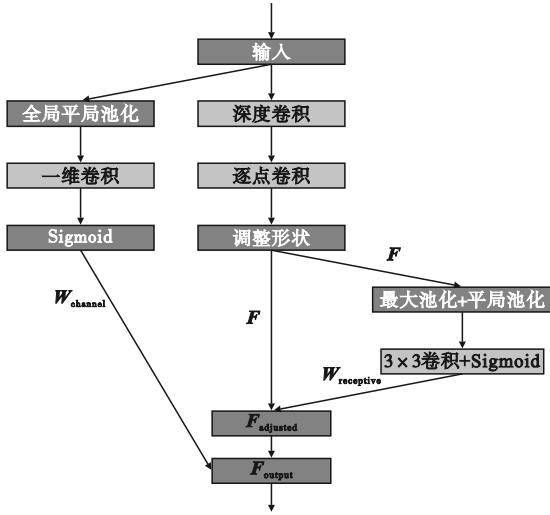


图2 ARFConv模块结构

Fig. 2 Structure of ARFConv module

首先,ARFConv模块通过深度可分离卷积来减少计算复杂度,并有效提取输入特征图中的多层次信息。其次,输入特征图通过全局平均池化生成通道注意力权重,这些权重通过一维卷积与Sigmoid激活函数进行非线性变换,生成自适应的通道加权系数,用于选择和调整每个通道的重要性。通道注意力机制通过自适应地加权每个通道的特征,使网络能够专注于对肿瘤检测至关重要的特征。通过这一机制,网络不仅能够自动选择并增强对肿瘤特征贡献较大的通道,还能有效抑制背景和无关区域的过度响应,从而提高检测精度。随后,ARFConv模块通过最大池化和平均池化对获得的多尺度特征图 F 提取显著信息,并利用 3×3 卷积和Sigmoid激活函数生成感受野权重。计算过程如下。

$$W_{receptive} = \text{Sigmoid}(\text{Conv}3 \times 3(\text{MaxPool}(F) + \text{AvgPool}(F))) \quad (1)$$

式中: F 表示多尺度特征图;MaxPool表示全局最大池化;AvgPool表示全局平均池化; $W_{receptive}$ 表示生成的感受野权重。权重 $W_{receptive}$ 进一步与多尺度特征图 F 逐元素相乘,调整每个位置的感受野响应,表达式为

$$F_{adjusted} = F \odot W_{receptive} \quad (2)$$

式中: $F_{adjusted}$ 为调整后的特征图; \odot 表示逐元素相乘。

最后,生成的通道权重 $W_{channel}$ 用于自适应地

调整每个通道的注意力权重,与调整后的特征图 $F_{adjusted}$ 逐元素相乘,获得输出 F_{output} ,表达式为

$$F_{output} = F_{adjusted} \odot W_{channel} \quad (3)$$

与现有模块相比,ARFConv模块的通道注意力机制通过全局平均池化生成权重以优化通道特征选择,通过引入动态感受野调整,不仅实现了多尺度特征的精确提取,还突破了仅对固定感受野加权的限制。上述设计使得ARFConv在复杂背景和不规则目标形态的检测任务中更加鲁棒。

1.2 分割融合卷积模块

SFCM旨在解决传统卷积在多尺度特征捕捉和复杂背景处理中的局限性。现有方法在处理不规则病灶特征时难以精准提取不同尺度和方向上的特征信息。为此,SFCM通过多尺度特征提取与融合策略,提升了模型对复杂背景和不规则特征的表达能力。

输入特征首先通过CBS调整通道数并重组特征,为后续处理提供灵活而高效的特征表达。随后,SFCM设计了一系列多尺度特征提取策略,初始特征图经过三次最大池化操作,生成多尺度的特征图,通过分层处理输入特征,增强对上下文信息的感知能力。上述池化结果与初始特征图在通道维度上拼接,形成一个信息更完整的特征图。

针对传统卷积固定感受野的不足,SFCM引入了特征分割与深度卷积增强模块。输入特征图被分为两部分,一部分保留原始特征,另一部分进入深度卷积增强模块,通过水平方向卷积(1×3 ,膨胀率2)、垂直方向卷积(3×1 ,膨胀率2)以及 1×1 卷积融合特征。增强后的特征与原始特征逐元素相乘,实现自适应融合,其计算公式为

$$X_{output} = X \odot ((X * K_h) * K_v * K_{(1 \times 1)}) \quad (4)$$

式中: X 为输入特征; X_{output} 为输出特征; K_v 为水平方向的深度膨胀卷积核,卷积核大小为 1×3 ,膨胀率为2; K_h 为垂直方向的深度膨胀卷积核,卷积核大小为 3×1 ,膨胀率为2; $K_{1 \times 1}$ 是用于特征融合和压缩的 1×1 卷积核;*表示卷积操作。

增强后的特征图与未处理的特征图再次在通道维度上拼接,生成更为多样和丰富的特征表示。通过 1×1 卷积和CBS组合卷积层进一步融合与压缩特征图,得到最终输出结果。

1.3 可变形融合模块

传统卷积操作由于固定的采样位置和卷积核形状,难以准确捕捉复杂的肿瘤特征,尤其在应对不规则形状的脑肿瘤或早期变异肿瘤检测时,其局限性更为明显。本文将DCNv2引入YOLOv8

中的 C2f 模块中^[9]。DCNv2 通过引入动态学习偏移量的机制,使卷积核能够根据输入图像的几何特征自动调整位置,从而更好地适应脑肿瘤的形态变化。具体来说,DCNv2 在每个卷积采样点引入可学习的偏移量,使得感受野不再局限于规则的矩形,而是能够自适应地调整以捕获脑肿瘤图像的实际形状。

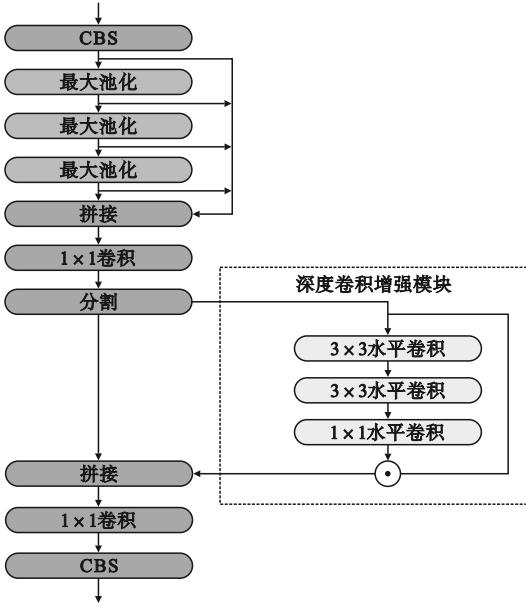


图 3 SFCM 结构图

Fig.3 Structural diagram of SFCM

DCNv2 的核心在于引入动态学习偏移量的机制,使卷积核能够根据输入图像的几何特征自动调整位置。通过在每次卷积操作中学习一个偏移值,卷积核不仅可以在标准位置采样,还能自适应地关注不同形状和位置的肿瘤区域。DCNv2 卷积操作表达式为

$$y(p_0) = \sum_k w_k \cdot x(p_0 + p_k + \Delta p_k) \quad (5)$$

式中: y 为输出值; p_0 为输出位置; k 为采样点索引; w_k 为卷积核的权重; x 为输入特征值函数; p_k 为卷积核的标准采样位置; Δp_k 为学习得到的偏移量。由此,卷积操作不仅仅是基于固定位置的采样,而是自适应调整以捕获更有意义的局部特征,增强了模型在处理复杂病变区域时的灵活性。

本文在改进算法设计中,将 DCNv2 与 YOLOv8 的 C2f 模块中 Bottleneck 结构进行融合,提出了 DFB 结构。DFB 在减少模块计算量的同时,显著提升了网络对几何变形的建模能力,增强了对不规则形状脑肿瘤的检测性能。将原始 Bottleneck 结构的第一个卷积层替换为 DCNv2,赋予模型更强的形态建模能力。该结构的具体设计如图 4 所示。

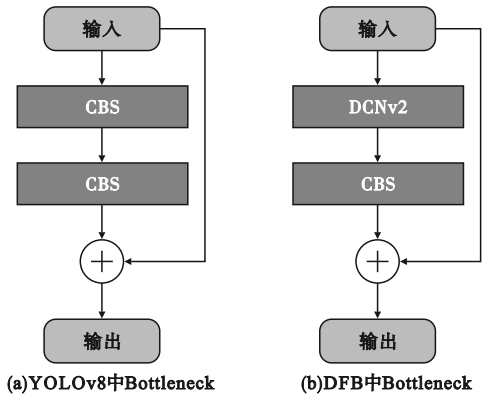


图 4 YOLOv8 中 Bottleneck 与 DFB 中的 Bottleneck

Fig.4 Bottleneck in YOLOv8 and DFB

2 实验分析与讨论

2.1 实验数据集

为了评估提出模型的性能,本研究使用了 Br35H 数据集,图像样例如图 5 所示。数据集包含 701 张图像,划分为训练集和测试集两个部分,其中 500 张图像用于训练集,201 张图像用于测试集。输入图像的分辨率为 640×640 像素。

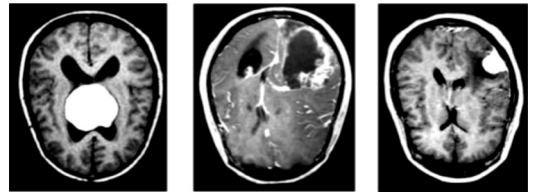


图 5 数据样例

Fig.5 Data example

2.2 实验配置与评估指标

在模型训练和推理过程中,操作系统使用 Ubuntu 20.04 LTS,处理器为 AMD EPYC 9754 128-Core,深度学习框架采用 PyTorch 1.11.0,Python 版本为 3.8。图形处理器使用 NVIDIA RTX 4090D。训练过程设置为 100 个训练轮次。

为了全面评估模型在脑肿瘤检测任务中的性能,本文选择了以下评估指标:精确率 (Precision)、召回率 (Recall)、平均精度均值 (mAP@0.5、mAP@0.5:0.9) 和计算量 (GFLOPs),以全面反映模型在不同检测任务中的表现,帮助有效比较模型的检测效果和计算效率。

2.3 TumorNet-YOLO 模型实验结果

本文提出的 TumorNet-YOLO 模型在脑肿瘤 MRI 图像检测任务中表现出优异性能。实验结果显示,模型 Precision 为 98.2%,Recall 为 89.7%,

mAP@0.5 达到 96.6%, mAP@0.5:0.9 为 73.8%。

图 6 展示了模型在 mAP@0.5 下的性能变化趋势,随着训练迭代次数增加,mAP@0.5 稳步上升并最终达到 96.6%,表明模型在 IoU 阈值为 0.5 的情况下具备优良的目标检测能力和收敛性。此外,图 7 展示了模型在 mAP@0.5:0.9 下的表现。该指标综合不同 IoU 阈值下的检测性能,最终达到 73.8%,证明模型在多种复杂场景下仍能保持较高的检测精度。图 6 和图 7 表明,模型在低 IoU 和高 IoU 阈值下均展现出卓越的检测性能,体现了其在应对多样化目标形态和重叠区域时的强适应能力。图 8 展示了模型的预测结果,其中检测框标注了置信度分数,结果表明模型能够准确定位脑肿瘤区域,大多数检测框置信度较高,接近真实标签。然而,在某些复杂场景中,置信度偏低(如 0.4),可能受到图像质量或背景干扰的影响。尽管如此,模型在多种肿瘤形态和位置的检测中仍展现出较高的稳定性和准确性。

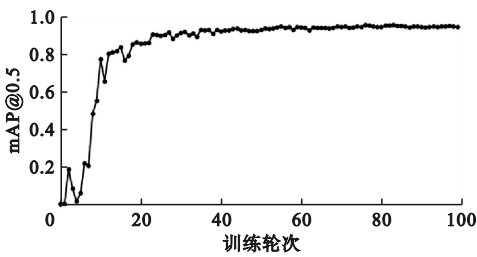


图 6 mAP@0.5 曲线图

Fig. 6 mAP@0.5 plot

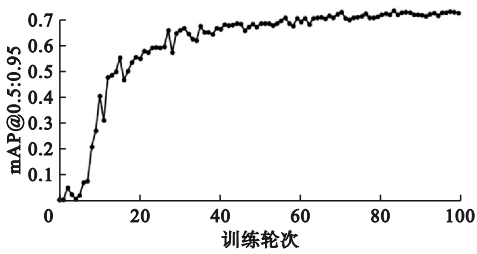


图 7 mAP@0.5:0.9 曲线图

Fig. 7 mAP@0.5:0.9 plot

2.4 不同模型对比实验

为验证本文模型的有效性,本文选择多个主流检测模型,包括 YOLOv3^[10]、YOLOv5^[11]、YOLOv7^[12]、YOLOv8^[13]、YOLOv9^[14]、YOLOv10^[15]、RT-DETR^[16]和 RCS-YOLO,与本文提出的模型进行对比,评估各模型在不同指标上的性能表现。

对比结果如表 1 所示,本文模型 Precision 达到 98.2%,较基础 YOLOv8 提升了 5.2%,表明模型能够有效降低误检率,其高精确率得益于引入

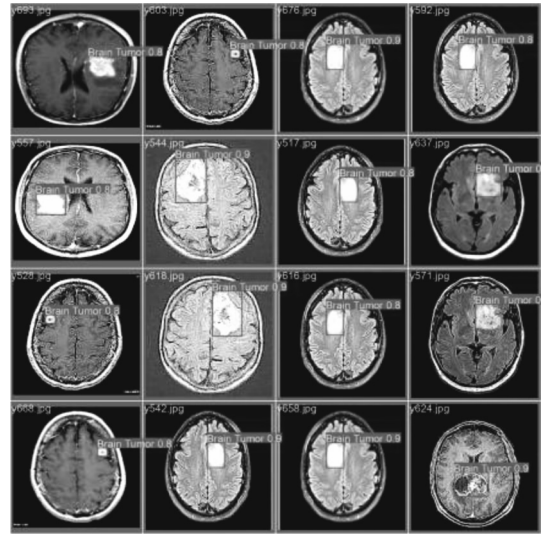


图 8 预测结果

Fig. 8 Predicted results

ARFConv 和 SFCM 模块的特征提取优化。在计算复杂度方面,本文模型计算量 GFLOPs 仅需 8.6,低于复杂模型如 RT-DETR-L 和 RCS-YOLO,展现出优越的性能与效率平衡。该性能优化主要归因于 DFB 模块的轻量化设计和高效卷积策略的应用。较低的计算开销使模型能够满足实时检测需求,但在进一步降低复杂度的同时保持高检测精度仍是未来的优化方向。

表 1 对比实验结果

Table 1 Comparison of experimental results

模型	Precision / %	Recall / %	mAP@0.5 / %	mAP@0.5:0.9 / %	GFLOPs
YOLOv3-L	91.6	90.5	93.3	70.3	282.2
YOLOv3-T	91.5	89.6	92.8	69.7	18.9
YOLOv5-n	93.8	88.7	95.6	69.1	7.1
YOLOv7-n	92.0	91.8	93.4	69.7	13.0
YOLOv8-n	93.0	92.2	94.5	70.5	8.1
YOLOv9-n	95.3	87.1	94.7	72.1	7.6
YOLOv10-n	93.5	84.1	92.7	68.6	6.5
RT-DETR-L	94.9	91.1	93.4	69.3	103.4
RCS-YOLO	94.4	92.5	94.8	71.3	105.2
本文模型	98.2	89.7	96.6	73.8	8.6

2.5 DCNv2 与其他注意力机制的对比实验

为验证本文将 DCNv2 引入 YOLOv8 中 C2f 模块对模型性能提升的有效性,本文选择了 SK-Net^[17]、CA^[18]和 CBAM^[19]机制替换 DCNv2 进行对比实验,结果如表 2 所示。实验结果表明,本文模型在 mAP@0.5 和 mAP@0.5:0.9 两项指标上

均取得了最佳表现,充分展示了 DCNv2 与 C2f 模块结合在提升模型性能方面的显著优势。

表 2 不同模块对比

Table 2 Comparison of different modules %

替换模块	Precision	Recall	mAP@0.5	mAP@0.5:0.9
SKNet	93.6	90.5	95.2	71.0
CA	93.0	90.0	95.6	71.5
CBAM	94.3	91.7	96.3	72.2
本文模型	98.2	89.7	96.6	73.8

2.6 泛化实验

为了验证本文提出模型的通用性,在 BTOD 和 BCCD 两个不同的数据集上进行对比实验。BTOD 是一个公开的脑肿瘤检测数据集,包含 385 张 MRI 图像,肿瘤区域通过手动标注的边界框标注,BTOD 数据集可在 Kaggle 平台获取,包含 364 张血液涂片图像,涵盖三类不同的细胞^[20]。泛化实验结果如表 3 所示。

表 3 泛化实验结果

Table 3 The results of the generalization experiments %

数据集	模型	Precision	Recall	mAP@0.5	mAP@0.5:0.9
BTOD	YOLOv8n	87.2	84.4	92.1	61.4
	本文模型	90.0	88.8	95.5	69.8
BCCD	YOLOv8n	86.0	89.9	92.6	63.5
	本文模型	87.5	90.1	93.4	64.8

从表 3 可以看出,本文提出的模型在多个数据集上的表现均优于基线模型 YOLOv8n。在 BTOD 数据集上,本文模型的 mAP@0.5 和 mAP@0.5:0.9 分别为 95.5% 和 69.8%。在 BCCD 数据集上,本文模型的 mAP@0.5 和 mAP@0.5:0.9 分别提升至 93.4% 和 64.8%。上述结果表明本文模型能够在多种数据集上提供更好的检测性能。

3 结论

本文提出的改进模型在脑肿瘤 MRI 图像分析中表现优异,通过自适应感受野卷积、分割融合卷积和可变形融合模块,提升了复杂背景和不规则肿瘤特征的处理能力。实验表明,模型在检测精确率和平均精度上均优于传统 YOLO 系列及主流检测模型,特别在复杂肿瘤检测方面表现突

出。尽管如此,仍需优化召回率以减少复杂肿瘤区域的漏检问题。此外,在高分辨率 MRI 或实时检测中,模型计算的效率存在瓶颈,需进一步简化结构。

参考文献 (References):

- [1] RANJBARZADEH R, BAGHERIAN K A, JAFARZADEH G S, et al. Brain tumor segmentation based on deep learning and an attention mechanism using MRI multi-modalities brain images[J]. Scientific Reports, 2021, 11(1): 10930.
- [2] CHATTOPADHYAY A, MAITRA M. MRI-based brain tumour image detection using CNN based deep learning method [J]. Neuroscience Informatics, 2022, 2(4): 100060.
- [3] TAHIR B, IQBAL S, USMAN GHANI KHAN M, et al. Feature enhancement framework for brain tumor segmentation and classification[J]. Microscopy Research and Technique, 2019, 82(6): 803–811.
- [4] ALMUFAREH M F, IMRAN M, KHAN A, et al. Automated brain tumor segmentation and classification in MRI using YOLO-based deep learning[J]. IEEE Access, 2024, 12: 16189–16207.
- [5] KANG M, TING C M, TING F F, et al. RCS-YOLO: a fast and high-accuracy object detector for brain tumor detection [C]//Medical Image Computing and Computer Assisted Intervention-MICCAI 2023. Cham, Switzerland: Springer, 2023: 600–610.
- [6] WAGEH M, AMIN K, ALGARNI A D, et al. Brain tumor detection based on deep features concatenation and machine learning classifiers with genetic selection [J]. IEEE Access, 2024, 12: 114923–114939.
- [7] CHANU M M, SINGH N H, MUPPALA C, et al. RETRACTED ARTICLE: computer-aided detection and classification of brain tumor using YOLOv3 and deep learning[J]. Soft Computing, 2023, 27(14): 9927–9940.
- [8] GU W C, SUN K X. AYOLOv5: improved YOLOv5 based on attention mechanism for blood cell detection[J]. Biomedical Signal Processing and Control, 2024, 88: 105034.
- [9] YU J F, SUN L, SONG S H, et al. BAIDet: remote sensing image object detector based on background and angle information[J]. Signal, Image and Video Processing, 2024, 18(12): 9295–9304.
- [10] TANG Y L, JIN S H, BIAN G, et al. Shipwreck target recognition in side-scan sonar images by improved YOLOv3 model based on transfer learning[J]. IEEE Access, 2020, 8: 173450–173460.
- [11] ZHONG S H, CHEN C L, LUO W S, et al. DDVC-YOLOv5: an improved YOLOv5 model for road defect detection[J]. IEEE Access, 2024, 12: 134008–134019.
- [12] CUI W Q, LI Z Y, DUANMU A N, et al. CCG-YOLOv7: a wood defect detection model for small targets using improved YOLOv7[J]. IEEE Access, 2024, 12: 10575–10585.
- [13] NIU S Q, XU X L, LIANG A, et al. Research on a lightweight method for maize seed quality detection based on improved YOLOv8[J]. IEEE Access, 2024, 12: 32927–32937.
- [14] CHIEN C T, JU R Y, CHOU K Y, et al. YOLOv9 for fracture detection in pediatric wrist trauma X-ray images[J]. Electronics Letters, 2024, 60(11): e13248.